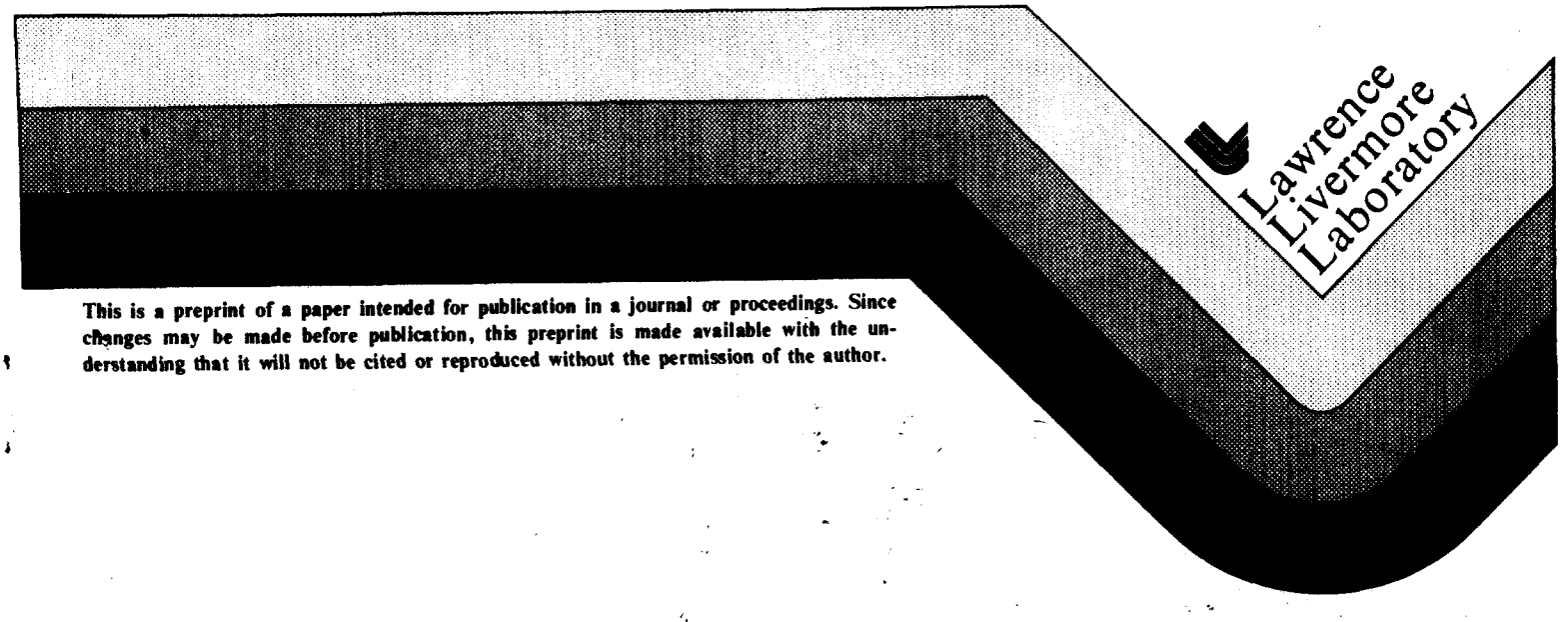
UCRL- 84437
PREPRINT

PARAMETER ESTIMATION FROM NOISY TRANSIENT ELECTROMAGNETIC MEASUREMENTS

D. T. Gavel
J. V. Candy
D. L. Lager

This paper was prepared for submittal to
Nuclear EMP Meeting
Anaheim, California

August 5-7, 1980



PARAMETER ESTIMATION FROM NOISY TRANSIENT ELECTROMAGNETIC MEASUREMENTS

D. T. Gavel, J. V. Candy, D. L. Lager
Lawrence Livermore Laboratory, Livermore, CA 94550

Abstract

The study of EMP phenomenon has promoted the development of techniques to investigate transient electromagnetic response data. Characterization of the EMP transient response information is necessary to evaluate the performance of that system in a hostile environment. An efficient technique to characterize this performance is to fit an electromagnetic model to the data.

In this paper we describe the performance of three different signal processing techniques applied to parameterize a body from noisy experimental electromagnetic transient response data. We briefly describe the techniques which range from the well-known Prony method to the more sophisticated extended Kalman filter and finally to the highly sophisticated maximum likelihood identifier. We compare the performance of these algorithms and discuss their tradeoffs.

1.0 Introduction

The study of EMP phenomenon has promoted the development of techniques to investigate transient electromagnetic response data. The characterization of EMP transient response information is a matter of national concern. Since large amounts of data are necessary to pointwise define an arbitrary transient response, it is quite reasonable to "identify" a parameterization or model of the "response". The model developed is useful, not only to merely

Work performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore Laboratory under Contract No. W-7405-ENG-48.

parameterize the response, but also to give more meaningful information about the physical process producing the response itself.

In this paper we discuss the implementation of signal processing algorithms which can be used to "estimate" the parameters of an electromagnetic response model from noisy transient measurements. The techniques employed range from simplified algorithms which perform well for high signal-to-noise ratios, to complex model-based estimators, which perform well for low signal-to-noise ratios. In Section 2 we present the necessary background information. The various algorithms are discussed (simply) in Section 3. In Section 4, the application to transient EM data is presented.

2.0 Background

In electromagnetic wave theory it is possible to represent the response of an object to various excitations by the singularity expansion method (SEM). The SEM represents an electromagnetic variable (field, current, etc.) as the impulse response of the object [1], i.e.,

$$\underline{U}_p(\underline{r},t) = \sum_i \eta_i(\underline{e},s_i)\underline{v}_i(\underline{r})e^{s_i t} \qquad (1)$$

where

- $\underline{U}_p$ vector impulse response
- η complex coupling coefficient
- $\underline{v}$ complex natural mode describing the behavior of $\underline{U}_p$ over the object
- $\underline{e}$ exciting field characteristics (e.g., polarization, direction of evidence, etc.)
- $\underline{r}$ spatial coordinates or position
- s_i complex natural frequency (or pole, or natural resonance)

The sets of parameters ($\{s_i\}$, $\{\underline{v}_i(\underline{r})\}$) are dependent on the object parameters only and independent of the excitation. The effect of the

exciting wave is contained entirely within the set of coupling coefficients $\{\eta_i(\underline{e}, s_i)\}$ which are independent of the position on the body. Thus, the electromagnetic interaction is completely characterized by these sets. In fact, the response to an arbitrarily shaped waveform can be generated using concepts of linear system theory where the response y to an arbitrary impulse is given by the convolution(*)

$$y(\underline{r},t) = U_p(\underline{r},t) * u(\underline{r},t) \tag{2}$$

Implicit in (2) is that $\underline{e}$ is the same for the new exciting waveform. However, $(\{s_i\}, \{\underline{v}_i(\underline{r})\})$ are invariant; therefore, these sets can be used to "parameterize" a given object for <u>any</u> excitation. We are not concerned at this point in the partitioning of the natural modes and coupling coefficients, so we define the set of complex residues at a point $\underline{r}_o$ as

$$\underline{C}_i(\underline{r}_o) := \eta_i(\underline{e},s_i)\underline{v}_i(\underline{r}_o) \tag{3}$$

and for this work concern ourselves only with scalar response functions.† Thus, the impulse response of the linear system of (2) can be represented as

$$y(t) = H(t) * \delta(t) = \sum_{i=1}^{N} C_i e^{s_i t} \tag{4}$$

where

$s_i := \sigma_i + j\omega_i$, σ the damping ratio and ω the natural frequency

---

†It should be noted that the sophisticated model-based estimators discussed subsequently can be used to identify separately the η_i and $\underline{v}_i$ parameters if desired as well as vector response functions (multiple measurement instruments); however, this work was not feasible in the allotted time.

H(t) is the object impulse response at position $\underline{r}_o$, i.e., $U_p(\underline{r}_o, t)$. Thus, the parameterization of the object can be stated simply by the electromagnetic parameter estimation problem.

"Given a set of noisy electromagnetic response† measurements $\{z(t)\}$, find the "best" (minimum variance) estimates ($\{\hat{\sigma}_i, \hat{\omega}_i\}$, $\{\hat{C}_i\}$) characterizing an unknown object."

We will assume that the noise contaminates the response y as

$$z(t) = y(t) + v(t) \tag{5}$$

where v is zero mean, Gaussian with covariance R.

Before we begin discussing the various estimation algorithms applied to this problem, we must define an alternate way of representing a linear system which is equivalent to (3) and (4). Recall from ordinary differential equations [2] that (4) represents the solution of a N^{th} order differential equation. It is well known that this equation can be broken down to the solution of N first order differential equations of the general form:

$$\dot{\underline{x}}(t) = F\underline{x}(t) + \underline{g}\, u(t)$$

$$y(t) = \underline{h}^T\underline{x}(t) \tag{6}$$

where

x is the N-state vector, u, y are the respective input and output.
F is a N x N matrix and g, h are N-vectors††

---

†This representation is not limited only to scalar systems, e.g. u and y can be vectors and g, h^T become matrices.

††This representation is not limited only to scalar systems, e.g. u and y can be vectors and g, h^T become matrices.

This representation is called the "state space" form in linear system theory [2] and forms the basis of various parameter estimation schemes (e.g., see [4]). It is easily shown that the impulse response of (6) is

$$y(t) = \underline{h}^T e^{Ft} \underline{g} = \sum_{i=1}^{N} C_i e^{s_i t} \tag{7}$$

or in transfer function form, we have

$$H(s) = \underline{h}^T (sI-F)^{-1} \underline{g} = \sum_{i=1}^{N} \frac{C_i}{(s+s_i)} \tag{8}$$

In the next section, we discuss three parameter estimation algorithms applied to this problem: (1) Interactive Prony's technique which utilizes the models of (6), or (8); (2) extended Kalman filter technique; and (3) the maximum likelihood identifier, both of which use the state space form of (6).

3.0 Parameter Estimation Algorithms

In this section we discuss the three parameter estimation algorithms employed to extract the set of object parameters ($\{\hat{\sigma}_i, \hat{\omega}_i\}$, $\{\hat{C}_i\}$) from noisy measurement data. The algorithms employed were: (i) Interactive Prony's technique (IPT); (ii) extended Kalman filter (EKF); and (iii) maximum likelihood identifier (MXLKID). We will not discuss the mathematical details of these algorithms, but rather include the primary references for the interested reader (see Appendices A, B, and C for details). After presenting each algorithm, we will compare them and discuss the various tradeoffs.

The Interactive Prony's technique (IPT) is basically a linear least squares estimator for poles in the discrete (z transform) domain [1]. The algorithm is depicted (simply) in Fig. 1. Depending on the signal-to-noise ratio (SNR) [3] either the impulse response (high SNR) or the autocorrelation response (low SNR) is estimated using fast Fourier transforms or sample autocorrelation estimators, respectively. The filtered data is then "windowed" and poles estimated from each data window solving a set of linear

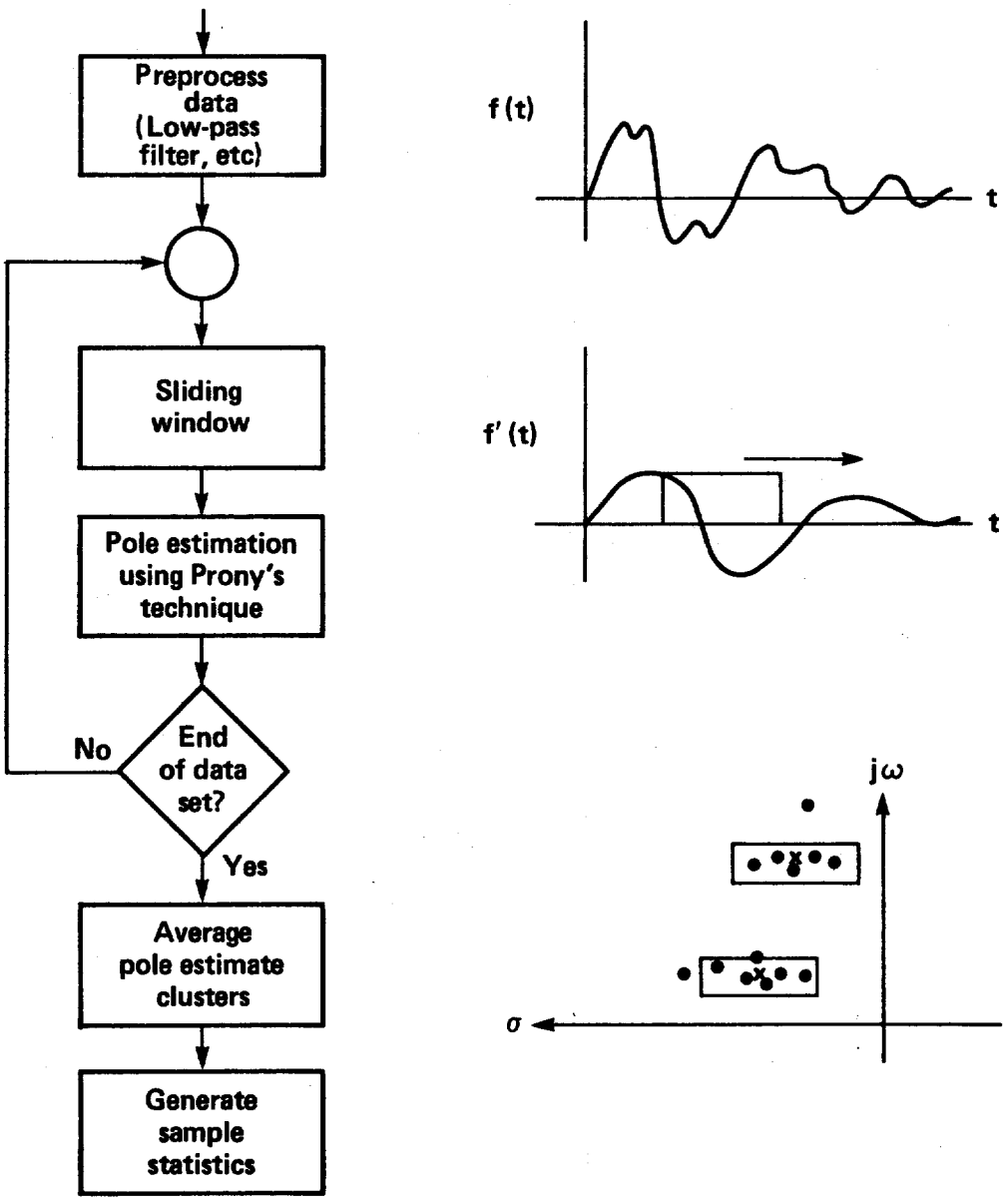


Fig. 1 Interactive Prony Technique

matrix equations to obtain linear least squares estimates of the coefficients of a polynomial, the roots of which are the discrete (z domain) poles. These poles are then transformed to the continuous domain and identified directly with the object response [1]. This technique is repeated by the processor many times and "pole clusters" are obtained. The average pole values are then selected and sample statistics calculated. It should be noted that the discrete or sampled data domain representation is necessary because of the use of "sampled" response data. Not accounting for the sampling phenomenon, will result in erroneous estimates for the continuous poles.

The extended Kalman filter (EKF) is basically a nonlinear state estimation algorithm which can be used to estimate unknown parameters by redefining them as states. Recall that a state estimator is a computer algorithm which may incorporate: (i) knowledge of the physical process phenomonology (EM response); (ii) knowledge of the measurement system; (iii) knowledge of process and measurement uncertainties in the form of mathematical models to produce an estimate of the state.

Most state estimators can be placed in a recursive form with the various subtleties emerging in the calculation of the current estimate ($\hat{x}_{old}$). The standard tecnnique employed is based on updating the current estimate as a new piece of measurement data becomes available. The state estimates generally take the recurrence form

$$\hat{x}_{new} = \hat{x}_{old} + K_k \varepsilon_{new} \tag{9}$$

where

$$\varepsilon_{new} = z_k - \hat{z}_{old} = z_k - h(\hat{x}_{old}) \tag{10}$$

Here we see that the new state estimate is obtained by correcting the old estimate by a K-weighted amount. The term ε_{new} is the new information or

innovations[4], i.e., it is the difference between the actual measurement and the predicted measurement ($\hat{z}_{old}$) based on our old state estimate. The computation of the weight K depends on the error criterion used (e.g., mean-squared, absolute, etc.) [5, 6, 7].

Note that a physical process model (e.g. state equation in (6) for linear case) is used to produce $\hat{x}_{old}$. The interested reader should see Gelb[6] for details.

Thus, the EKF is a state estimator capable of producing estimates for nonlinear as well as linear processes and measurements. A simplified diagram of the algorithm is depicted in Fig. 2. Here we see that the state estimate $\hat{x}_{old}$ is calculated (or predicted) based on the process model, after the estimator is initialized. The calculation of the gain, K, and innovations, ε, follows. Note that the measurement at a given time step is utilized in calculating the current ε. From these calculations the new or corrected state estimate is obtained. The algorithm continues in this loop processing masurement data as it becomes available. This processing is considered on-line because it can be accomplished in conjunction with the response measurements, i.e., the state estimates are updated in real time, each time a new measurement becomes available.

The final algorithm is the maximum likelihood identifier (MXLKID). The MXLKID algorithm is a complex off-line technique which utilizes a parameter optimization algorithm looped around the EKF to obtain parameter estimates. The algorithm maximizes the likelihood function, or equivalently minimizes the negative log-likelihood function $J(\underline{\theta})$, i.e.

$$\min_{\theta} J(\underline{\theta}) = -1/2 \ell n(2\pi) - 1/2 \sum_{i=1}^{N} \varepsilon_{new}^{T}(i,\underline{\theta}) \left(R_{\varepsilon}(i,\underline{\theta})\right)^{-1} \varepsilon_{new}(i,\theta) +$$

$$\ell n \left| R_{\varepsilon}(i,\underline{\theta}) \right| \qquad (11)$$

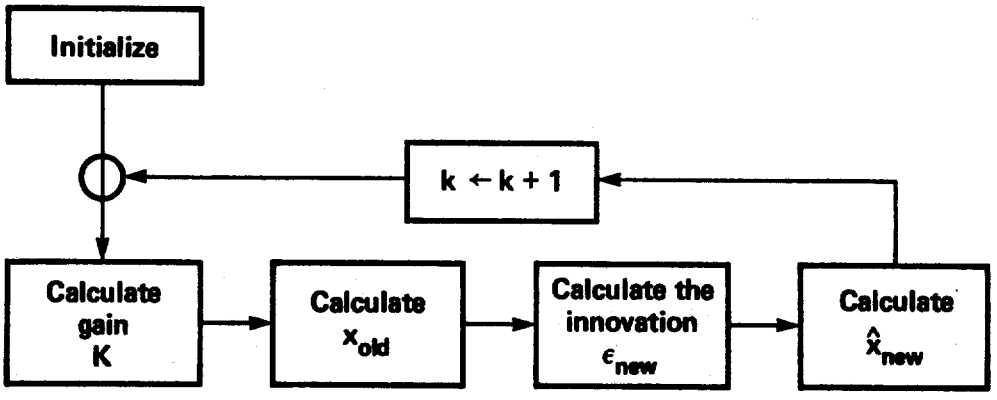


Fig. 2 Extended Kalman Filter Algorithm

where

ε_{new} is the innovation of equation (10)

R_{ε} is the corresponding innovation covariance matrix.

The parameter estimator (optimization algorithm) is usually a gradient-based technique [8, 9], i.e.,

$$\hat{\theta}_{new} = \hat{\theta}_{old} + \rho H \frac{\partial J}{\partial \underline{\theta}} \qquad (12)$$

where $\hat{\theta}$ are the parameter estimates

ρ is a step parameter

H is a weighting matrix dependent on the particular optimization technique used).

The simplified algorithm operation is depicted in Fig. 3. The Kalman filter is used to produce uncorrelated innovations, ε, from the correlated measurements, z. The log-likelihood function (11) is calculated using results from the Kalman filter. In some optimization algorithms the filter is also used to calculate elements in the weighting matrix, H (e.g. see [9]).

Before we discuss the performance of these three algorithms on the electromagnetic parameter estimation problem, we first compare their basic attributes. Referring to Table I, we see that the interactive Prony algorithm is a simple technique valid for high SNR, and because of the lack of system modeling it is restricted in scope of application (linear, time invariant problems only). Of course, because of its simplicity, it is less complex and faster than the EKF and MXLKID algorithms. The IPT requires many runs to generate an ensemble of samples for statistical validation of the parameter estimates whereas the two other techniques have statistical validation built in.

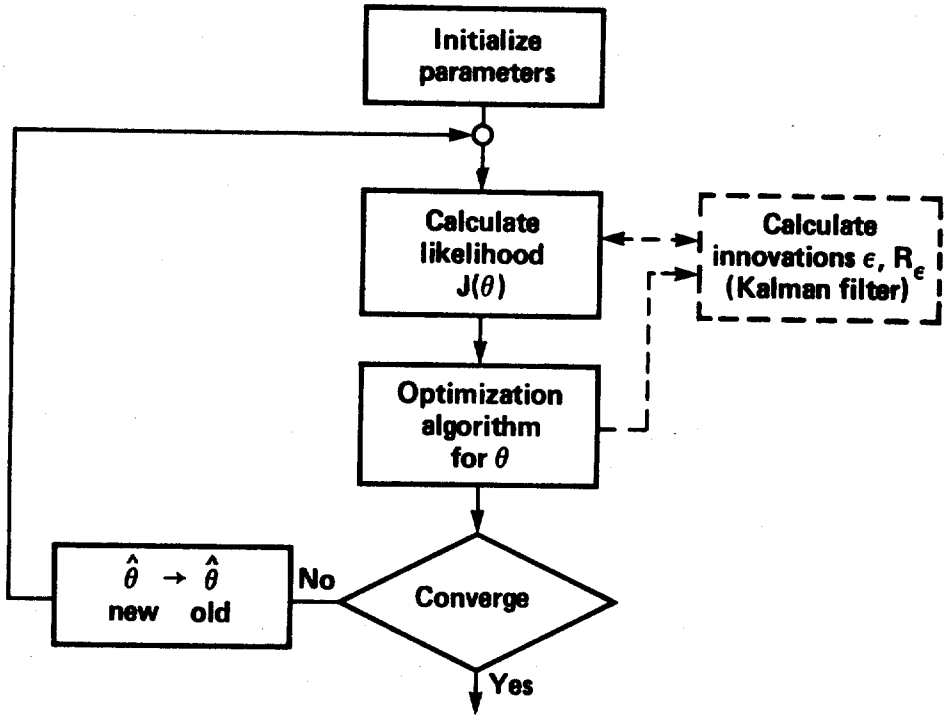


Fig. 3 Maximum Likelihood Identification Algorithm

TABLE 1: Identification Algorithm Comparisons

	IPT	EKF	MXLKID
Problem Scope	linear, time invariant	linear, nonlinear, time varying	linear, nonlinear
Signal Models	scalar	vector	vector
Noise Models	none	stationary, non-stationary	stationary, non-stationary
Complexity	simple†	complex	very complex
Application	off-line	on-line	off-line
Limitations	nigh SNR	medium SNR	low SNR
Computer Time	small†	medium	large
Accuracy	reasonable	reasonable	excellent
Statistical Validation	sample calculations†	generated	generated

† Neglecting FFT, Autocorrelation, and/or Ensemble Statistic Calculations

The EKF and MXLKID algorithms appear similar in many categories, which is expected since the MXLKID technique actually uses an EKF as an integral part of its computation. The main differences between MXLKID and the EKF are application, complexity, and accuracy. The MXLKID algoritnm is more accurate, however the price paid is complexity, computer time and the necessity of running the algorithm off-line, i.e., with a complete set of measurements available beforehand. If enough data is available the EKF can yield comparable results.

This completes our discussion of the parameter estimation algorithms. In the next section we discuss the application of these techniques to the EM parameterization problem.

4.0 EM Transient Response Application

The three parameter estimation techniques presented in the previous sections have been applied to transient electromagnetic pulse (EMP) response data. We begin this section with a brief description of the experiment which produced the data of interest, then discuss the approach taken to abstract the pole parameters. The initial phase of the pole extraction approach is the acquisition and preliminary analysis of the data. Certain assumptions concerning signal and noise modeling must be made before parameter estimation techniques can be applied. In the last part of this section we discuss the signal and noise modeling assumptions made, then we present pole extraction results for the three sets of data analyzed.

The experimental configuraton (Fig. 4) consists of a monocone source antenna, a movable verticle ground plane, and a taret cylinder all mounted on

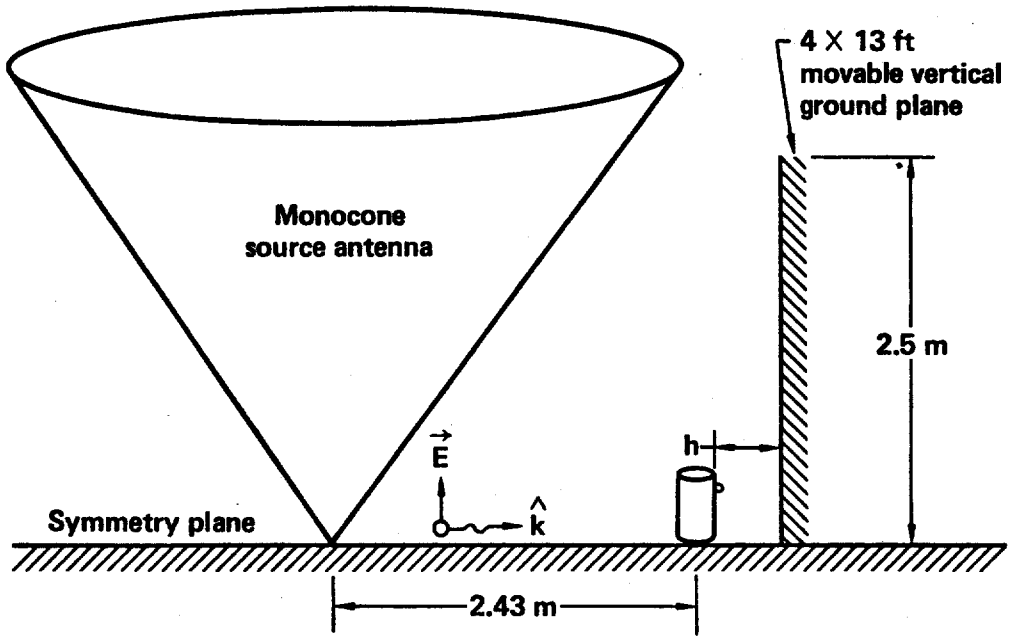


Fig. 4 Experimental Configuration for the
EM Transient Experiment

a conducting symmetry plane. A sensor for detecting current flows is located near the base of the cylinder as shown in Fig. 5. The sensor placement allows observation of the odd-narmonic response modes which are induced on the cylinder and its image in the symmetry plane.

The response of the cylinder to an electromagnetic pulse is dependent on the distance, h, from the axis of the cylinder to the reflecting ground plane. We acquired data for three such responses, as shown in Fig. 6, corresponding to h=∞ (ground plane removed), h=50 cm, and h=10 cm. More resonant responses are attained as h becomes smaller.

The signal modeling approacn is the Singularity Expansion Method (SEM) introduced earlier [1]. The electromagnetic variable (current in this case) is represented as a sum of singular response as in equation (1), which we repeat below:

$$\underline{U}_p(\underline{r},t) = \sum_i \eta_i(\underline{e},s_i)\underline{v}_i(\underline{r})e^{s_i t} \qquad (13)$$

since the probe location and exiting field characteristics are fixed for a given experiment we can lump $\underline{r}$ and $\underline{e}$ dependent terms into one complex residue which we designate C_i. The measurement process then extracts the real part of the response and also includes random measurement noise, $\underline{v}$. The parameter model for this experiment with measurement noise is the following:

$$z(t) = \mathrm{Re}[U_p(t)] + v(t) = \sum_{i=1}^{N} [C_i e^{s_i t}] + v(t) \qquad (14)$$

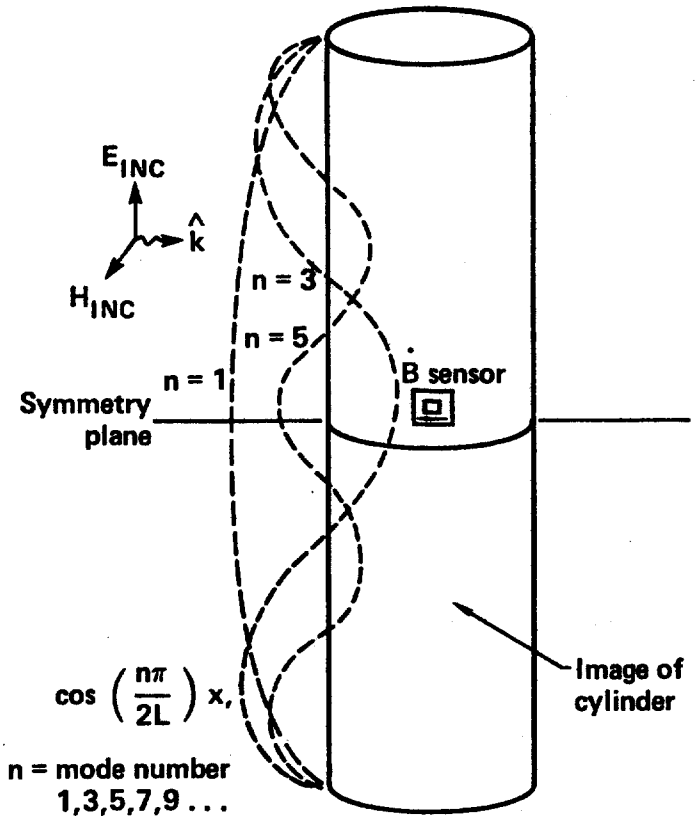


Fig. 5 Current Modes on a Cylinder over a Symmetry Plane

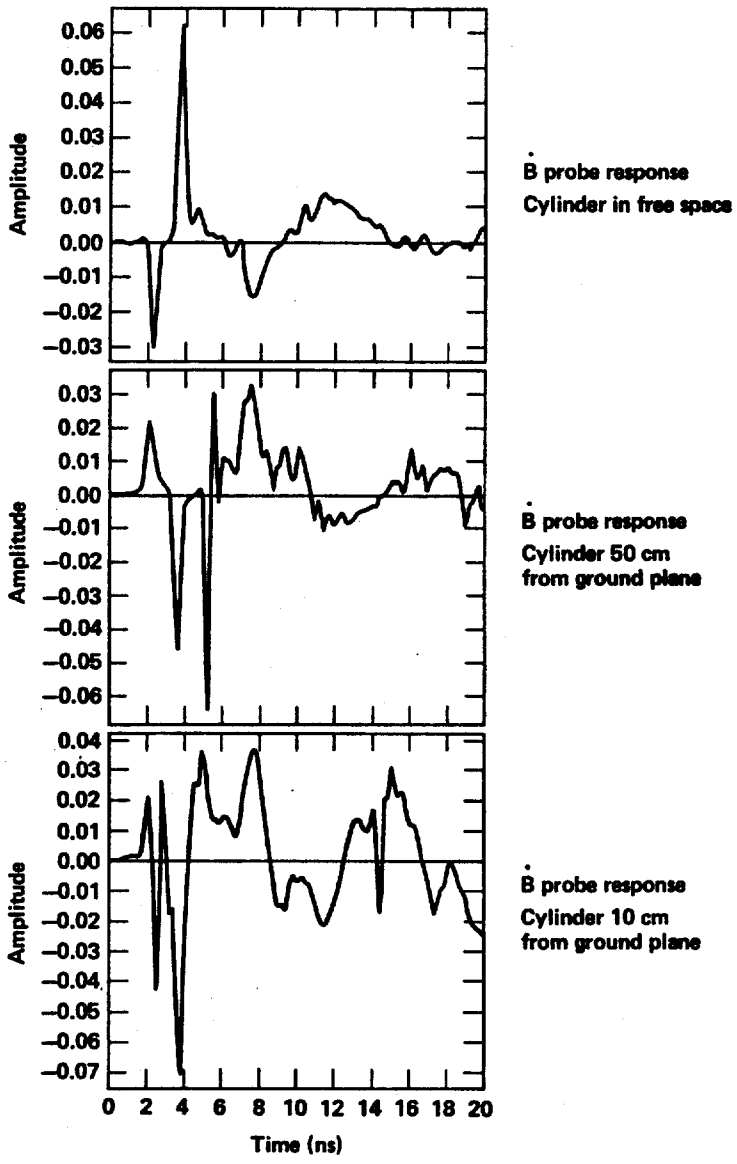


Fig. 6 Data was Acquired from Three Separate EM Transient Experiments

where

$z(t)$ is the measured signal of interest

$v(t)$ is the measurement noise

$N = 6$*

The pole extraction problem then becomes:

$$\text{Identify } (C_i, \omega_i, \sigma_i) \text{ for } i = 1,2,3 \text{ given } \{z(t)\}$$

An alternative noise model considers process as well as measurement noise. Process noise is correlated (as opposed to independent or "white" noise characteristic of a measurement probe) and typically has power spectral components within the signal bandwidth of interest. Process noise sources can include unmodeled EMP reflections, unknown environmental electromagnetics, and signal mismodeling. Process noises are modeled as driving noises, and hence are convolved with the impulse response of the object.

The following parametric model includes both process and measurement noises:

$$z(t) = \sum_{i=1}^{N} C_i e^{s_i t} + \int_{\tau=0}^{t} w_i(\tau) e^{s_i (t-\tau)} d\tau + v(t) \qquad (15)$$

where

$w_i(t)$ is process driving noise

* We chose the first three harmonics (N=6, i.e., a complex conjugate pair of poles for each harmonic) to form our signal model, however, because of probe location, very little of the second harmonic was observable, therefore the forthcoming pole extraction results will apply to the fundamental and third harmonic only.

The pole extraction problem in now:

$$\text{Identify } \{C_i, \omega_i, \sigma_i, \text{Cov}(w_i)\}$$
$$\text{for } i = 1,2,3 \text{ given } \{z(t)\}$$

We have the additional task of identifying the covariance (noise power) of the driving noise. To extract poles from the three response data sets shown in Fig. 6, the Prony technique, extended Kalman filter (EKF) and two versions of the maximum likelihood identifier were applied. The Prony technique accounts only for a measurement noise model, whereas the EKF uses both a process and a measurement noise model. The maximum likelihood identifier was designed for both cases. Version A accounted only for measurement noise whereas Version B included both process and measurement noise models.

The results of the pole extraction are summarized in Figs. 7 through 12*. Fig. 13 shows the migration of the fundamental pole as distance between the cylinder and ground plane becomes smaller. The results from each of the data analysis methods are compared to predicted results from an analytical study [9]. Fig. 14 shows a similar plot for the third harmonic.

We can make the following conclusions concerning the application of signal processing/parameter identification techniques to EM transient response analysis. In this paper we have shown that a variety of signal processing techniques are available for pole extraction each having trade offs in the degree of complexity, accuracy, and the scope of signal and noise models which can be used. After applying pole extraction techniques to real EM transient data we can generally conclude from our results that data with a higher

* The tabulated results for EKF and maximum likelihood techniques show 95% confidence intervals for the estimated parameter values. These confidence intervals are themselves estimates generated by the processing algorithm based upon the single data set being processed. The confidence intervals should not be confused with results from a thorough statistical analysis of the experiment, which would involve processing an ensemble (a large statistical set) of data records.

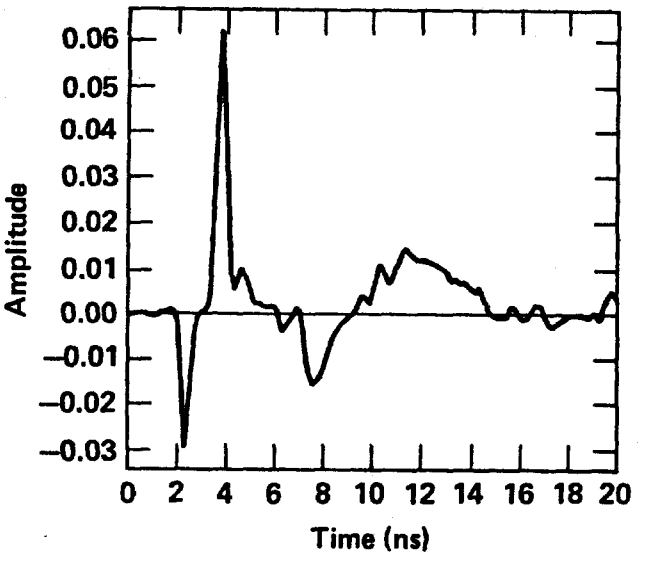
signal-to-noise ratio (i.e., the more resonant responses) provide more consistent pole estimates. The pole extraction has also, to some degree independently verified analytical predictions [9] for this experimental configuration.

5.0 Summary

In this paper we have discussed and compared the tradeoffs of three popular signal processing algorithms - the interactive Prony technique, the extended Kalman filter, and the maximum likelihood identifier and applied them to an electromagnetic transient response experiment. The results indicate that more consistent estimates were obtained with high signal-to-noise ratio signals. The results also independently validate theoretical prediction.

Acknowledgements

We would like to thank Mr. F. Deadrick and J. Zicker for data collection and performing the interactive prony processing. We would also like to thank Ms. Lopez for typing this manuscript.

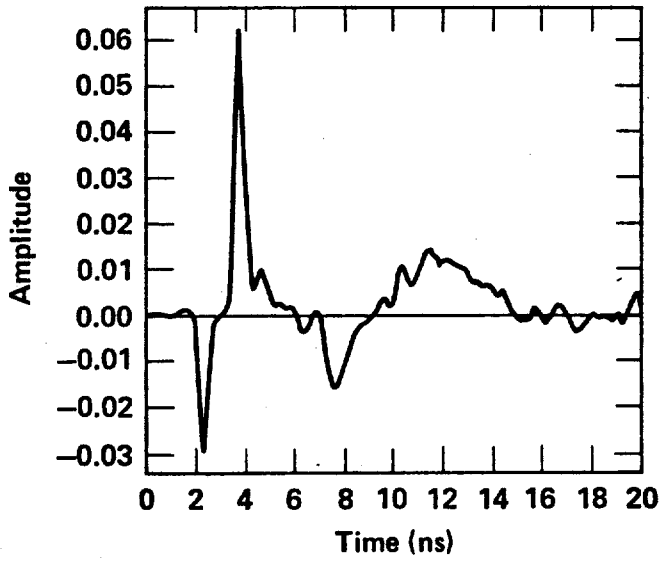


		Analytical prediction*	Interactive Prony algorithm	Extended Kalman filter	Maximum likelihood A+	Maximum likelihood B+
Pole no. 1	σ	−131.4	−99.6	−190 ± 38	−215 ± 62	−137 ± 235
(Fundamental)	ω	775.4	720.7	750 ± 9	526 ± 49	829 ± 154

*T. H. Shumpert and D. J. Galloway, "Transient Analysis of a Finite Length Cylindrical Scatterer Very Near a Perfectly Conducting Ground," Sensor and Simulation Notes, Note 226, August 1976.

+A: with measurement noise model only
B: with both measurement and process noise models included

Fig. 7 Pole Extraction Results - Free Space, Fundamental Pole

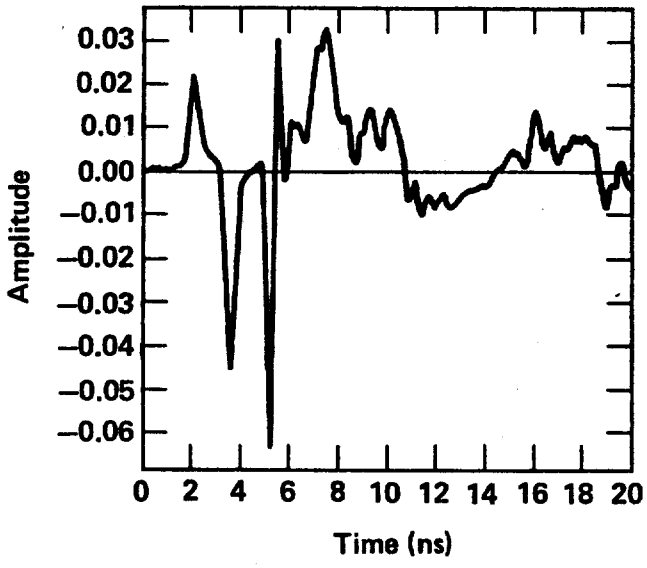


		Analytical prediction*	Interactive Prony algorithm	Extended Kalman filter	Maximum likelihood A+	Maximum likelihood B+
Pole no. 2 (3rd harmonic)	σ	−236	−297.3	−710 ± 36	−721 ± 451	−654 ± 693
	ω	2523.6	2433.5	2250 ± 27	1578 ± 147	2480 ± 308

*T. H. Shumpert and D. J. Galloway, "Transient Analysis of a Finite Length Cylindrical Scatterer Very Near a Perfectly Conducting Ground," Sensor and Simulation Notes, Note 226, August 1976.

+A: with measurement noise model only
B: with both measurement and process noise models included

Fig. 8 Pole Extraction Results - Free Space, Third Harmonic

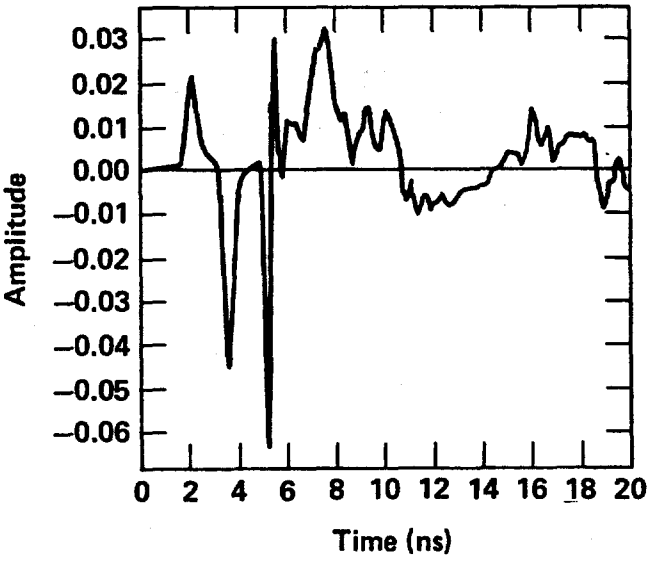


		Analytical prediction*	Interactive Prony algorithm	Extended Kalman filter	Maximum likelihood A+	Maximum likelihood B+
Pole no. 1 (Fundamental)	σ	−66.3	−74	−65 ± 31	−102 ± 46	−70 ± 127
	ω	736.2	541	736 ± 8	770 ± 32	722 ± 72

*T. H. Shumpert and D. J. Galloway, "Transient Analysis of a Finite Length Cylindrical Scatterer Very Near a Perfectly Conducting Ground," Sensor and Simulation Notes, Note 226, August 1976.

+A: with measurement noise model only
B: with both measurement and process noise models included

Fig. 9 Pole Extraction Results - h=50 cm
Fundamental Pole

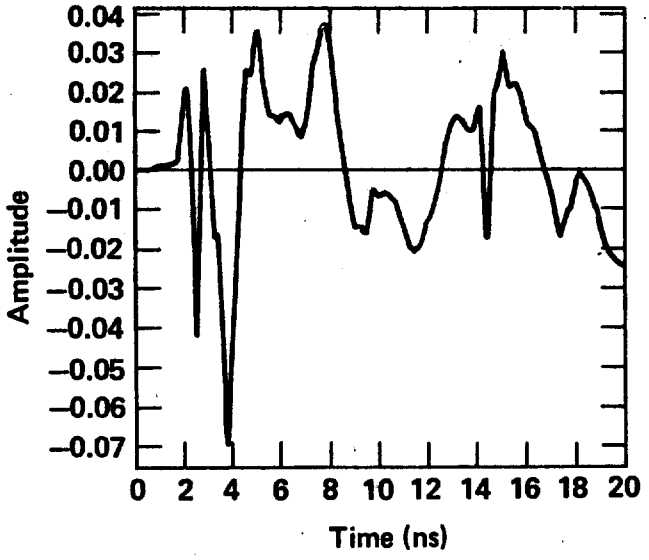


		Analytical prediction*	Interactive Prony algorithm	Extended Kalman filter	Maximum likelihood A+	Maximum likelihood B+
Pole no. 2 (3rd harmonic)	σ	Not available	−258.9	−460 ± 36	−440 ± 220	−206 ± 311
	ω		2383.2	2208 ± 24	2310 ± 96	2166 ± 144

*T. H. Shumpert and D. J. Galloway, "Transient Analysis of a Finite Length Cylindrical Scatterer Very Near a Perfectly Conducting Ground," Sensor and Simulation Notes, Note 226, August 1976.

+A: with measurement noise model only
B: with both measurement and process noise models included

Fig. 10 Pole Extraction Results - h=50 cm, Third Harmonic

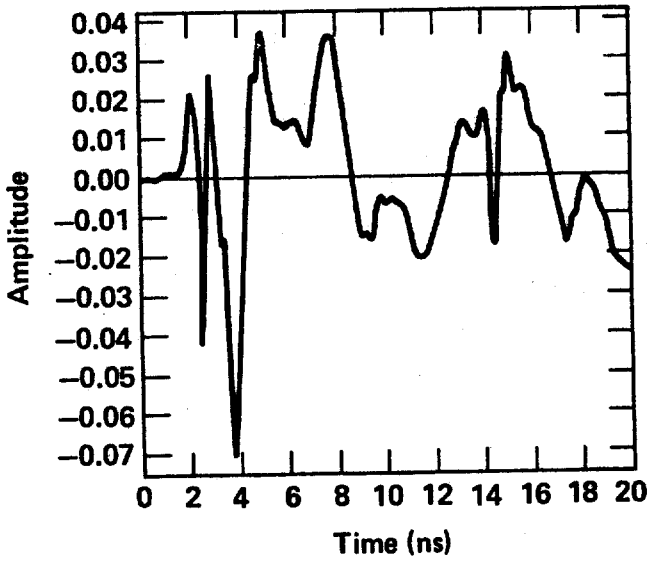


		Analytical prediction*	Interactive Prony algorithm	Extended Kalman filter	Maximum likelihood A+	Maximum likelihood B+
Pole no. 1	σ	−16.9	−22.9	−25 ± 23	−28.9 ± 13	−25 ± 49
(Fundamental)	ω	824	661	758 ± 6	767 ± 9	775 ± 15

*T. H. Shumpert and D. J. Galloway, "Transient Analysis of a Finite Length Cylindrical Scatterer Very Near a Perfectly Conducting Ground," Sensor and Simulation Notes, Note 226, August 1976.

+A: with measurement noise model only
B: with both measurement and process noise models included

Fig. 11 Pole Extraction Results - h=10 cm
Fundamental Pole



		Analytical prediction*	Interactive Prony algorithm	Extended Kalman filter	Maximum likelihood A+	Maximum likelihood B+
Pole no. 2 (3rd harmonic)	$\frac{\sigma}{\omega}$	Not available	−63.7 2311.6	−160 ± 34 2274 ± 18	−50.5 ± 25 2301 ± 27	−200 ± 52 2379 ± 30

*T. H. Shumpert and D. J. Galloway, "Transient Analysis of a Finite Length Cylindrical Scatterer Very Near a Perfectly Conducting Ground," Sensor and Simulation Notes, Note 226, August 1976.

+A: with measurement noise model only
B: with both measurement and process noise models included

Fig. 12 Pole Extraction Results - h= 10 cm
Third Harmonic

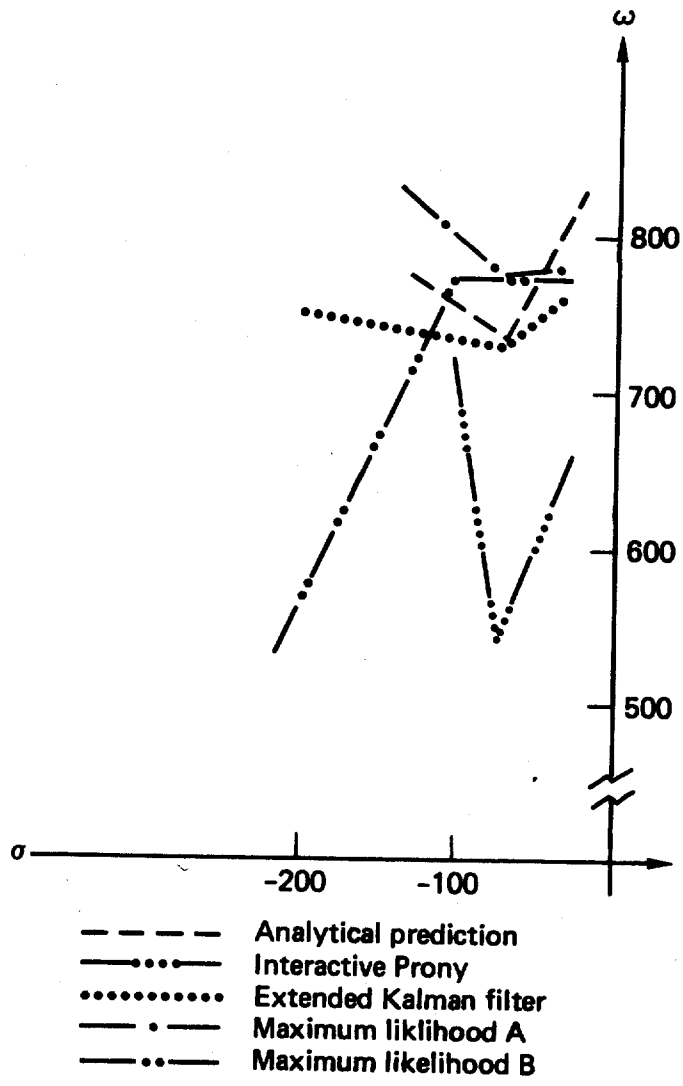


Fig. 13 Migration of the Fundamental Pole

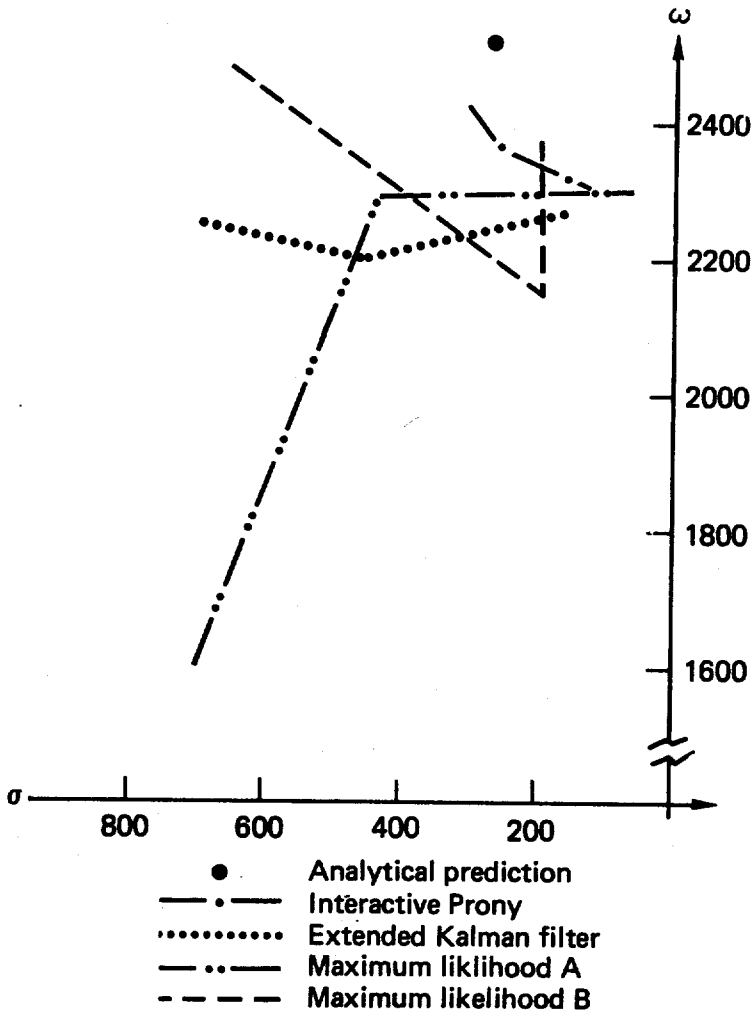


Fig. 14 Migration Of The Third Harmonic

References

[1] A. J. Poggio, M. L. Van Blaricum, E. K. Miller, and R. Mitra, "Evaluation of a Processing Technique for Transient Data," IEEE Trans. Antennas Propagat., AP-26, 1978.

[2] P. Padulo and M. Arbib, System Theory. Philadelphia: Saunders, 1974.

[3] W. D. Smith and D. L. Lager, "Parametric Characterization of Random Processes Using Prony's Method," Lawrence Livermore National Laboratory Report, UCRL-52673, 1979.

[4] T. Kailath, Lectures on Linear Least Squares Estimation, Berlin: Springer-Verlag, 1975.

[5] G. C. Goodwin and R. L. Payne, Dynamic System Identification, New York: Academic Press, 1977.

[6] A. Gelb, Applied Optimal Estimation, Boston: MIT Press, 1974.

[7] R. Castleton and J. V. Candy, "DYNEST, A Dynamic Estimator Calculation Program, Lawrence Livermore National Laboratory Report, UCRL-52573, 1978.

[8] N. K. Gupta and R. K. Mehra, "Computational Aspects of Maximum Likelihood Estimation," IEEE Trans. Autom. Contr., AC-19, 1974.

[9] D. T. Gavel, "MXLKID - A Maximum Likelihood Parameter Identifier," Lawrence Livermore National Laboratory Report, UCID-18744, 1980.

[10] T. H. Shumpert and D. J. Galloway, "Transient Analysis of a Finite Length Cylindrical Scatterer Very Near a Perfect Conducting Ground," Sensor and Simulation Notes, Note 226, August 1976.

Appendix A - Summary of properties of the time- and frequency-domain Prony techniques

	Time domain	Frequency domain
Defining equation	$f(t) = \sum_{\alpha=1}^{N} R_\alpha e^{s_\alpha t}$	$F(s) = \sum_{\alpha=1}^{N} R_\alpha/(s - s_\alpha)$
Rearrange equation (superscript denotes absence of $\beta = \alpha$ term)		$F(s) \prod_{\alpha=1}^{N} (s - s_\alpha) = \sum_{\alpha=1}^{N} R_\alpha \prod_{\beta=1}^{N}{}^{\alpha} (s - s_\beta)$
Introduce polynomial expansion	$\sum_{\alpha=0}^{N} a_\alpha X^\alpha$; $a_N = 1$	$\sum_{\alpha=0}^{N} a_\alpha s^\alpha$; $a_N = 1$
	$X = e^{s\delta}$	$\sum_{\beta=0}^{N-1} b_\beta^\alpha s^\beta$; $b_{N-1}^\alpha = 1$
Sample data	$f_i = f(i\delta)$; $i=0,\ldots,M_t$ (Note: equal-spaced samples)	$F_i = F(s_i)$; $i=0,\ldots,M_f$ (Note: samples can be arbitrarily spaced and located; $s_i = j\omega_i$ uses real-frequency sampling)
Develop linear system for polynomial coefficients	$\sum_{\alpha=0}^{N-1} M_{i\alpha} a_\alpha = -f_{N+i}$	$\sum_{\alpha=0}^{N-1} [M_{i\alpha} a_\alpha - N_{i\alpha} \tilde{b}_\alpha] = -s_i^N F_i$
	$M_{i\alpha} = f_{i+\alpha}$	$M_{i\alpha} = s_i^\alpha F_i$
	$i = 0,\ldots,M_t - N$	$N_{i\alpha} = s_i^\alpha$
		$i = 0,1,\ldots,M_f$
		$\tilde{b}_\beta = \sum_{\alpha=1}^{N} R_\alpha b_\beta^\alpha$
	(Note: system has N real unknowns and $M_t + 1 \geq 2N$)	(Note: system has 2N real unknowns and $M_f + 1 \geq N$)
Find poles s_α from	$\sum_{\beta=0}^{N} a_\beta X_\alpha^\beta = 0$	$\sum_{\beta=0}^{N} a_\beta s_\alpha^\beta = 0$
	$s_\alpha = \frac{1}{\delta} \ln X_\alpha$	$\alpha = 1,\ldots,N$
	$\alpha = 1,\ldots,N$	
Find residues R_α from	$\sum_{\alpha=1}^{N} M_{i\alpha} R_\alpha = f_i$	$\sum_{\alpha=1}^{N} M_{i\alpha} R_\alpha = F_i$
	$M_{i\alpha} = X_i^\alpha$	$M_{i\alpha} = 1/(s_i - s_\alpha)$
	$ii = 0,\ldots,N,\ldots,M_t$	$i = 0,\ldots,N,\ldots,M_f$
	where $M_t + 1 \geq 2N$	where $M_f + 1 \geq N$

Appendix B - Extended Kalman Filter Algorithm

Prediction:

$$\hat{x}_{k+1|k} = \hat{x}_{k|k} + \int_{k}^{k+1} f(\hat{x}_{\alpha|k}, u_{\alpha})\, d\alpha$$

$$\tilde{\Pi}_{k+1|k} = \int_{k}^{k+1} \left[F(\hat{x}_{k|k})\tilde{\Pi}_{\alpha|k} + \tilde{\Pi}_{\alpha|k}F^{T}(\hat{x}_{k|k}) + Q_{\alpha}\right] d\alpha$$

Innovation:

$$\varepsilon_{k+1} = z_{k+1} - h(\hat{x}_{k+1|k})$$

$$R^{\varepsilon}_{k+1|k} = H(\hat{x}_{k+1|k})\tilde{\Pi}_{k+1|k}H^{T}(\hat{x}_{k+1|k}) + R_{k+1}$$

Correction:

$$K_{k+1} = \tilde{\Pi}_{k+1|k}H^{T}(\hat{x}_{k+1|k})\,(R^{\varepsilon}_{k+1|k})^{-1}$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\varepsilon_{k+1}$$

$$\tilde{\Pi}_{k+1|k+1} = \left[I - K_{k+1}H(\hat{x}_{k+1|k})\right]\tilde{\Pi}_{k+1|k}\left[I - K_{k+1}H(\hat{x}_{k+1|k})\right]^{T} + K_{k+1}R_{k+1}K_{k+1}^{T}$$

where

$$F(\hat{x}_{k|k}) := \left.\frac{\partial f(x)}{\partial x}\right|_{x=\hat{x}_{k|k}}, \quad H(\hat{x}_{k|k}) = \left.\frac{\partial (h(x))}{\partial x}\right|_{x=\hat{x}_{k|k}}$$

and

$$\tilde{\Pi}_{k+1|k} := \mathrm{Cov}(\tilde{x}_{k+1|k}) \text{ for } \tilde{x}_{k+1|k} := x_{k+1} - \hat{x}_{k+1|k}$$

$$x_o \sim N(\hat{x}_{o|o}, \tilde{\Pi}_{o|o})$$

Appendix C - Maximum Likelihood Identifier Algorithm

Log Likelihood Function Calculation:

$$J(\underline{\theta}) = -1/2\ \ell n\ (2\pi) - 1/2 \sum_{i=1}^{N} \varepsilon^T(i,\hat{\underline{\theta}}_{OLD})\ (R^{\epsilon}(i,\hat{\underline{\theta}}_{OLD}))^{-1}\ \varepsilon(i,\hat{\underline{\theta}}_{OLD}) + \ell n \left| R^{\epsilon}(i,\hat{\underline{\theta}}_{OLD}) \right| *$$

Gradient Calculation:

$$\frac{\partial J}{\partial \underline{\theta}} := \left[\frac{\partial J(\theta)}{\partial \theta_{\ell}}\right] \approx \frac{J(\hat{\underline{\theta}}_{OLD} + \Delta\hat{\underline{\theta}}_{\ell}) - J(\hat{\underline{\theta}}_{OLD})}{\Delta\theta_{\ell}} \qquad \ell = 1, \ldots, p$$

Hessian Calculation:

$$\frac{\partial^2 J(\theta)}{\partial \underline{\theta}^2} := \left[\frac{\partial^2 J(\hat{\theta}_{OLD})}{\partial\hat{\theta}_j\ \partial\hat{\theta}_k}\right] = \sum_{i=1}^{N} \frac{\partial\varepsilon(i,\hat{\theta}_{OLD})}{\partial\hat{\theta}_j}\ (R_i^{\epsilon})^{-1}\ \frac{\partial\varepsilon(i,\hat{\theta}_{OLD})}{\partial\theta_k} +$$

$$+ 1/2\ \mathrm{tr}\left[(R_i^{\epsilon})^{-1}\ \frac{\partial R_i^{\epsilon}}{\partial\theta_j}\ (R_i^{\epsilon})^{-1}\ \frac{\partial R_i^{\epsilon}}{\partial\theta_k}\right]$$

$$+1/4\ \mathrm{tr}\left[(R_i^{\epsilon})^{-1}\ \frac{\partial R_i^{\epsilon}}{\partial\theta_j}\right] \mathrm{tr}\left[(R_i^{\epsilon})^{-1}\ \frac{\partial R_i^{\epsilon}}{\partial\theta_k}\right] \qquad (5)$$

* The quantities $\varepsilon(i,\underline{\theta})$ and $R^{\epsilon}(i,\underline{\theta})$, in these equations are generated recursively by the Kalman filter.

Parameter Estimate Update:

$$\underline{\hat{\theta}}_{NEW} = \underline{\hat{\theta}}_{OLD} + \rho \left[\frac{\partial 2 J(\theta)}{\partial \underline{\theta}} + \beta D^2 \right]^{-1} \frac{\partial J}{\partial \underline{\theta}}$$

Loop:

$\underline{\hat{\theta}}_{NEW} \rightarrow \underline{\theta}_{OLD}$

where

$J(\underline{\theta})$ is the scalar negative log-likelihood function

$\frac{\partial J}{\partial \underline{\theta}}$ is the p-gradient vector

$\frac{\partial^2 J(\theta)}{\partial \underline{\theta}^2}$ is the pxp Hessian matrix

$\underline{\theta}$ is the p-parameter vector

$\Delta\theta_\ell$ is the p-incremental parameter change vector (only ℓ^{th} element nonzero)

$\underline{\varepsilon}$ is the m-innovation vector

R^ϵ is the mxm innovation covariance matrix

D is the pxp diagonal Marquardt matrix (contains the square root of the diagonal elements of $\left(\frac{\partial^2 J(\theta)}{\partial \theta^2}\right)$

ρ is the scalar step adjustment

β is the Marquardt parameter used to weight the diagonal Marquardt Matrix